



685 Clyde Avenue
Mountain View, CA 94043
(650) 210-8600
Fax: (650) 210-9275

Fax

To: USPTO
Talivaldis Smits

From: Megan Carroll

Fax: (703) 746-6055

Pages: 2

Phone:

Date: August 6, 2003

Re: No 09/769,880

CC:

☐ **Urgent** ☐ **For Review** ☐ **Please Comment** ☐ **Please Reply** ☐ **Please Recycle**

Attached is our DRAFT response as we discussed by phone. Please call me when you have had a chance to review and we can schedule a follow up call with the inventor.

Thanks!

DRAFT/ FOR DISCUSSION PURPOSES ONLY

MEMO

TO: USPTO

Talivaldis Smits

RE: Response to Office Action

Application No. 09/769,880

First named inventor : Bertrand A. Damiba

FROM: BeVocal, Inc.

DATE: July 31, 2003

In our phone conversation of July 17, 2003 you indicated that you wanted clarification of what we meant by "tuning" and some clarification sufficient to enable a person skilled in the art to practice the invention. We hope the following information addresses your concerns.

What the application is intended to cover is a method and process for tuning speech application recognition. "Tuning", in general, is the process of making iterative and quantifiable improvements in recognition accuracy. These changes for the purposes of "tuning" may include making changes to the recognition grammar coverage; amending or altering the phonetic dictionaries; testing against multiple acoustic model sets; changing recognition engine parameters; changing endpointing parameters and the like.

The problem BeVocal addressed with this tuning platform was the efficient management and classification of huge amounts of data and the management of the transcription process in a way that would allow for the tuning of multiple, deployed, high-traffic applications. The method and process is a distributed method and may be web-based. It utilizes recorded live human interactions with the application in order to establish and subsequently improve the recognition accuracy.

Utterances and certain related log file information (such as the information referred to in Table 1) are automatically captured on production servers as part of the speech recognition process of the live application. The system allows for an administrator to determine which utterances are to be recorded. The system also can be set to determine which sets of utterances and log files are to be retrieved from production and copied over to a central database. The vocal transcriber program allows an administrator to set certain properties that determine which of those utterances or classes of utterances are to be transcribed. The program retrieves the appropriate data from the log files and updates the database with the selected information. The program also manages the transcription process, allocating the files to be transcribed among a pool of available human transcribers. Transcribers then can access the database, transcribe the utterances, certain information that may be associated with such utterances (such as the information referred

DRAFT/ FOR DISCUSSION PURPOSES ONLY

to in Table 1 and Table 3) and record other information such as discrepancies between the log file and what is actually heard, and store such information in the database.

The data can then be accessed by trained linguists who are able to easily run a variety of recognition accuracy tests because numerous types of information may be obtained through simple queries (such as the information referred to in Table 2 and Table 4), thus enabling iterative and quantifiable tuning changes that lead to improvements in recognition accuracy.